Synthesis of a New, Highly Fluorescent Amino Acid Derivative: *N*-[(*tert*-Butoxy)carbonyl]-3-[2-(1*H*-indol-3-yl)benzoxazol-5-yl]-L-alanine Methyl Ester

by **Katarzyna Guzow**, **Mariusz Szabelski**, **Joanna Malicka**, and **Wiesław Wiczk***

Faculty of Chemistry, University of Gdańsk Faculty of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 PL-Gdańsk (fax: (048)-(058)-342-03-57; e-mail: ww@chemik.chem.univ.gda.pl)

A simple method of synthesis of a new, highly fluorescent amino acid derivative from the simple and generally available substrates 3-nitro-L-tyrosine and 1*H*-indole-3-carbaldehyde is described. The obtained compound, *N*-[(*tert*-butoxy)carbonyl]-3-[2-(1*H*-indol-3-yl)benzoxazol-5-yl]-L-alanine methyl ester (**4**), possesses a high fluorescence quantum yield. The described method illustrates a new possibility of synthesis of amino acid derivatives possessing desirable photophysical properties.

Introduction. – Proteins contain three amino acid residues that contribute to their ultraviolet fluorescence: phenylalanine, tyrosine, and tryptophan. The emission of proteins is dominated by the tyrosine and tryptophan contributions, but mostly by that of tryptophan, which absorbs at the longest wavelength and displays the largest excitation coefficient. A complicating factor in the interpretation of protein fluorescence is the presence of multiple fluorescent amino acids in most proteins. Since the environment of each residue is distinct, the spectral properties of each residue are generally different. The emission spectra of tyrosine or tryptophan residues in proteins overlap at most usable wavelengths, and one cannot easily separate the spectral contributions of each tyrosine or tryptophan moiety in a multi-tyrosine or multi-tryptophan protein [1]. In addition, complex time-resolved fluorescence decays are found even for tryptophan itself [2][3] or simple tyrosine derivatives [4] as well as for proteins or peptides containing a single tyrosine or tryptophan residue [5–7]. For this reason, one cannot simply interpret a multi-exponential time-resolved decay in terms of the individual tyrosine or tryptophan residues in a protein. To some extent, the fluorescence-decay heterogeneity can be avoided by means of a constrained analogue of phenylalanine [8], tyrosine [9], or tryptophan [10][11]. The study of a particular protein or peptide in the presence of other tryptophan-containing proteins can be accomplished with tryptophan analogues that absorb at longer wavelengths than tryptophan. The most widely used tryptophan analogues are 5-hydroxytryptophan (5HW) [12][13] and 7-azatryptophan (7AW) [12][14–17]. These analogues display absorption at up to 320 nm, which is beyond the longest-wavelength absorption of tryptophan. The 5-hydroxytryptophan and 7-azatryptophan differ from each other and from tryptophan in their spectral properties [18]. In water, 5HW displays higher quantum yields than tryptophan (0.275 *vs.* 0.14). It is less sensitive to solvent polarity than tryptophan and displays an emission maximum near 339 nm. The 7AW is highly sensitive to solvent polarity, and its quantum yield decreases in contact with water. In

water, its quantum yield is low, near 0.017 with the emission maximum near 403 nm. A non-exponential fluorescence decay has been observed for 7AW and for peptides containing a 7AW residue [19]. Therefore, there is a need for a new non-proteinogenic amino acid that possesses better spectral properties than tryptophan or its above-mentioned analogues (absorption at longer wavelength, higher fluorescence quantum yield, and simple photokinetics). Available fluorescence probes [20][21] can be attached to a specific reactive group (amino, carboxyl, thiol) but cannot be incorporated in an optional place in the peptide chain because they do not possess an amino acid moiety. The 2-phenylbenzoxazoles are known as photostable, highly efficient UV dyes [22] and are used as organic brightening agents [23], laser dyes [22], and organic plastic scintillators [24]. Benzoxazoles are generally prepared by the action of mono- and dibasic acid derivatives on *o*-aminophenols, either by heating in a solvent or by thermal fusion without acid catalysis [25], in polyphosphoric acid [24][26] or by conversion of mixed *N*,*O*-diacylated 2-aminophenols [27][28]. There is also a method for synthesizing 2-substituted benzoxazoles by oxidative cyclization of azomethines (*Schiff* bases derived from aromatic 1-amino-2-hydroxy compounds). Lead tetraacetate [29], chloranil, *N*-bromosuccinimide (NBS), benzoyl peroxide [30], or barium manganate [31] are often used as the oxidizing agents.

Since the photophysical properties (long-wavelength absorption and high fluorescence quantum yield) of 2-arylbenzoxazoles are better than those of tryptophan and its analogues, and since these compounds may be prepared from generally available stubstrates (tyrosine and appropriate aromatic aldehydes), we decided to synthesize a derivative of [2-(1*H*-indol-3-yl)benzoxazol-5-yl]-L-alanine.

Results and Discussion. – *Synthesis. N*-[(*tert*-Butoxy)carbonyl]-3-[2-(1*H*-indol-3-yl)benzoxazol-5-yl]-L-alanine methyl ester (**4**) was synthesized from N^{α}-Boc-protected 3-amino-L-tyrosine methyl ester **2** as a substrate (Boc = (*tert*-butoxy)carbonyl), *via* the intermediate *Schiff* base, which underwent oxidative cyclization to the heterocyclic compound in the presence of lead tetraacetate or *N*-bromosuccinimide under mild conditions. Thus, *N*-[(*tert*-butoxy)carbonyl]-3-nitro-L-tyrosine methyl ester (**1**) was reduced to the corresponding-3-amino ester **2** by catalytic hydrogenation in MeOH (H_2, Pd/C). The *Schiff* base **3** was prepared according to the procedure described in [29] by addition of 1*H*-indole-3-carboxaldehyde in abs. EtOH to a solution of **2** in boiling abs. EtOH. Cyclization of the *Schiff* base **3** to the benzoxazole structure was performed in two ways, either with lead tetraacetate in AcOH [29] or with *N*-bromosuccinimide (NBS) in CH_2Cl_2 [30]. Both reactions gave the desired product **4** in *ca*. 10% yield. In the case of the cyclization with NBS a by-product, the bromo-substituted ester **5**, was isolated in *ca*. 4% yield.

The hydrogenation of the 3-nitro-L-tyrosine derivative **1** to the 3-amino compound **2** as well as the synthesis of the *Schiff* base **3** proceed with high yields, and both cyclization methods lead to the desired product **4**, although the yields of this latter step are rather low (not exceeding 10%), and this reaction step needs to be improved.

Photophysical Properties. Absorption spectra of the obtained compound **4** in MeOH, MeCN, and methylcyclohexane are presented in *Fig. 1*. In all solvents studied, the absorption maxima are at *ca*. 310–320 nm, and the absorbance decreases to zero at 360 nm. In the less-polar solvent methylcyclohexane, the vibronic structure is well-

Scheme

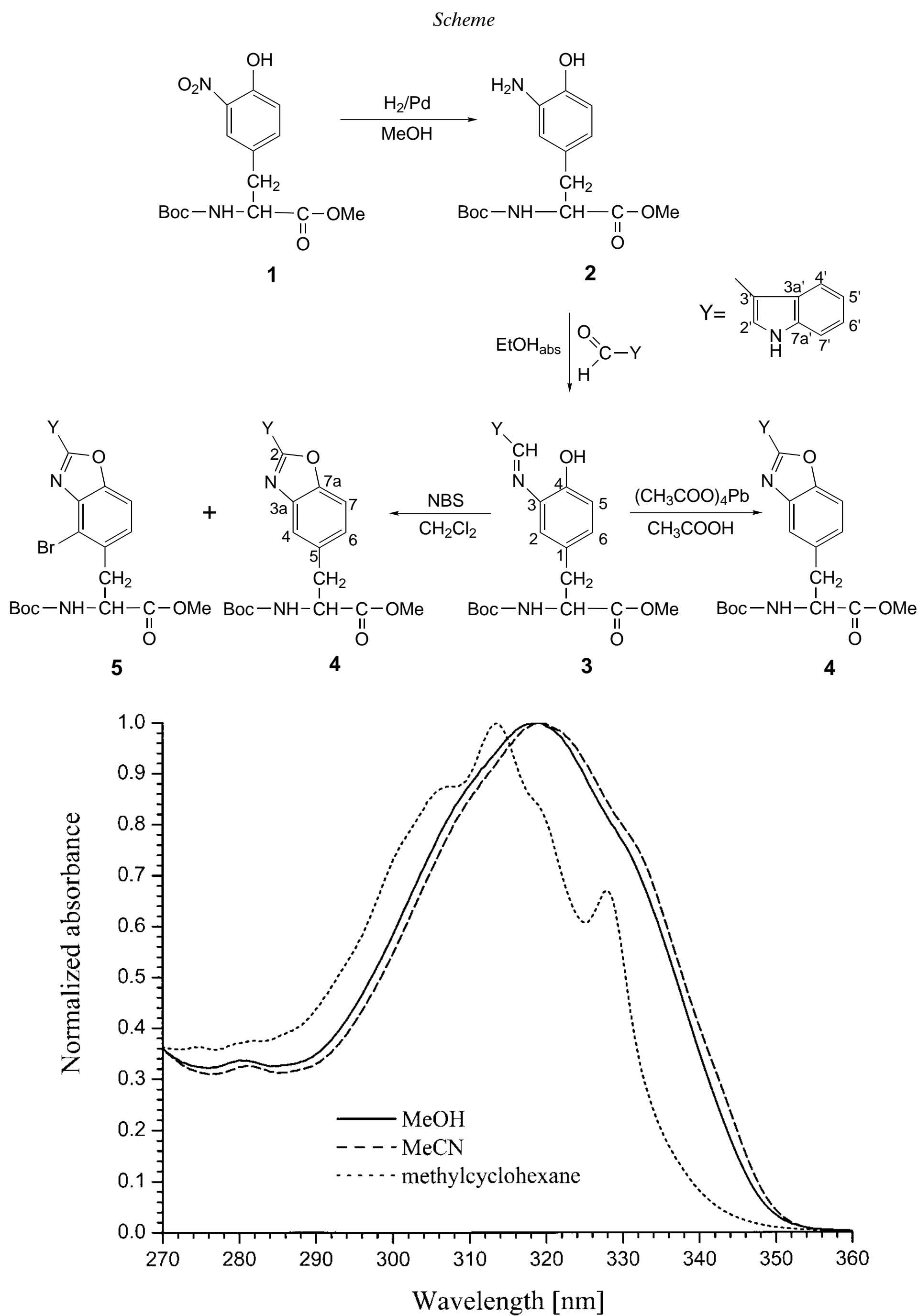

Fig. 1. *Absorption spectra of* N-*[(*tert-*butoxy)carbonyl]-3-[2-(1*H-*indol-3-yl)benzoxazol-5-yl]*-L-*alanine methyl ester* (**4**) *in methylcyclohexane, MeOH, and MeCN*

resolved, whereas in the more polar MeOH and MeCN, the absorption spectra are very similar with diffuse vibronic structure. The absorption band of **4** is shifted to longer wavelength compared to that of tryptophan. Also, the higher molar absorption coefficient of the (benzoxazol-5-yl)-L-alanine derivatives (*ca.* 20000 $dm^3 mol^{-1} cm^{-1}$) than that of tryptophan (*ca.* 5400 $dm^3 mol^{-1} cm^{-1}$) allows selective excitation of this compound in the presence of tryptophan. The absorption spectrum of the by-product **5** in all solvents studied is similar in shape but shifted to longer wavelength (by *ca.* 5 nm) compared to the absorption spectrum of the parent molecule **4**.

The bromo-substituted by-product **5** is non-fluorescent. The fluorescence spectra of the substituted benzoxazolyl-L-alanine methyl ester **4** (*Fig. 2*) show an emission maximum at 340 nm in methylcyclohexane and at 360 nm in MeOH and MeCN. Similar to the absorption spectrum, the emission spectrum of **4** in methylcyclohexane has a well-resolved vibronic structure, whereas in MeOH and MeCN, the fluorescence spectra exhibited a very poorly resolved vibrational structure. The values of the fluorescence quantum yields of the substituted benzoxazolyl-L-alanine derivative are high and slightly dependent on solvent polarity. In MeOH the fluorescence quantum yield θ is 0.77, in MeCN 0.82, and in methylcyclohexane 0.57. The high quantum yields and their increase with solvent polarity are consistent with data published by *Kanegae et al.* [22] for 2-phenylbenzoxazole derivatives without the amino acid moiety.

Conclusion. – The high fluorescence quantum yield of **4** and the high photostability of benzoxazole derivatives [22] make this amino acid derivative an interesting

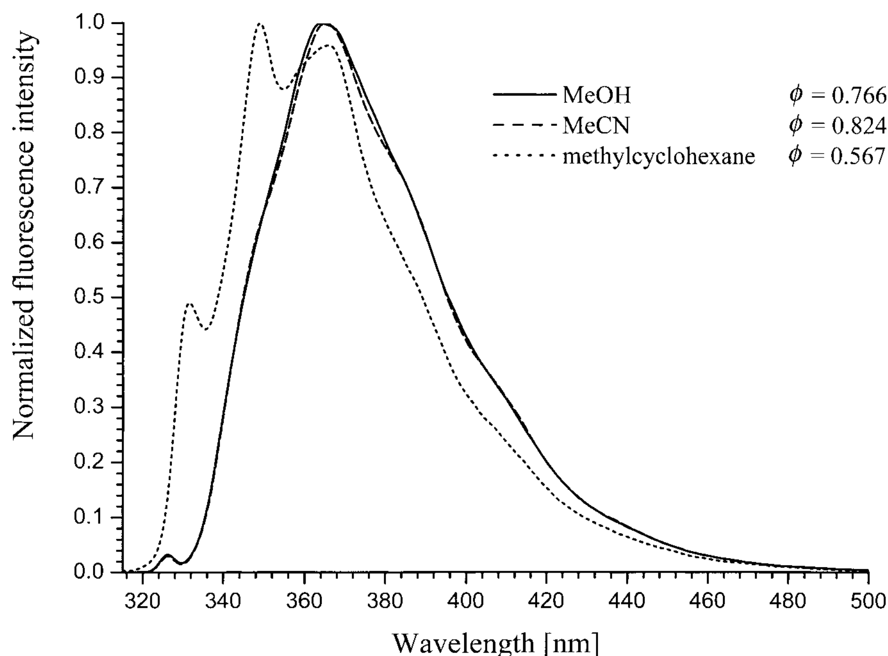

Fig. 2. *Emission spectra of* N-*[(*tert*-butoxy)carbonyl]-3-[2-(1*H*-indol-3-yl)benzoxazol-5-yl]*-L-*alanine methyl ester* (**4**) *in methylcyclohexane, MeOH, and MeCN*

alternative to the tryptophan analogues. This compound can be incorporated in a peptide chain and used as a convenient fluorescent probe. One of the disadvantage of amino acid derivative **4** is its bulkiness, which can cause a conformational change of a peptide backbone after its incorporation into a peptide chain. The probability of conformational changes decreases when the amino acid derivative **4** occurs at the N- or C-terminus of a peptide chain.

This work was supported by the *Polish State Committee for Scientific Research* (*KBN*) under grant BW-8000-5-0260-0.

Experimental Part

General. The 1*H*-indole-3-carboxaldehyde, lead tetraacetate, and NBS were purchased from *Lancaster*, and 3-nitro-L-tyrosine from *Fluka*. The 3-nitro-L-tyrosine methyl ester [32] and *N*-[(*tert*-butoxy)carbonyl]-3-nitro-L-tyrosine methyl ester [33] were prepared according to literature procedures. TLC: *Merck* silica gel *60* F_{254} plates; detection by UV (254 and 366 nm); solvent ratios in *v*/*v*. Semi-prep. reversed-phase HPLC: *Kromasil* column *C-8*, 250-mm long, inner diameter 20 mm, 5 μm. Anal. reversed-phase HPLC: *Kromasil* column C_{18}, 250-mm long, inner diameter 4.5 mm, 5 μm. M.p.: in capillary tubes; uncorrected. Optical rotation: *Perkin-Elmer* polarimeter, model *343*. Absorption and fluorescence spectra: *Perkin-Elmer Lambda-18* spectrophotometer and *Perkin-Elmer LS-50B* spectrofluorimeter, resp.; solvents MeOH, MeCN, and methylcyclohexane of either spectroscopic or HPLC grade; calculation of quantum yields θ with the reference tryptophan in water ($\theta = 0.14$ [34]) or quinine sulfate in 1N H_2SO_4 ($\theta = 0.546$ [35]). FT-IR: *Bruker IFS-66* spectrometer; KBr pellets; in cm^{-1}. ^1H-NMR, ^{13}C-NMR, $^1H,^1H$ TOCSY, $^1H,^{13}C$ HSQC: *Varian-Mercury 444-BB* spectrometer (400 MHz); δ in ppm, referenced internally to the residual proton resonance of $CDCl_3$ (=7.27 ppm) or to $SiMe_4$ (=0.00 ppm) for ^1H and to $CDCl_3$ (=77.23 ppm for centerline) for ^{13}C. FAB-MS: *Masslab Trio-3* spectrometer; in *m*/*z*. Elemental analysis: *Carlo Erba, Eager 200*.

*3-Amino-*N*-[(*tert*-butoxy)carbonyl]-*L*-tyrosine Methyl Ester* (**2**). *N*-[(*tert*-Butoxy)carbonyl]-3-nitro-L-tyrosine methyl ester (4 mmol, 1.36 g; m.p.90–91°; $[\alpha]_D^{20} = +6.21$ ($c = 0.1$M, MeOH)) was dissolved in MeOH (30 ml) and hydrogenated over Pd/C. After 90 min (TLC monitoring (CH_2Cl_2/MeOH/AcOH 100 : 10 : 1): R_f 0.9 (**1**), R_f 0.72 (**2**)), the catalyst and solvent were removed: **2** as white solid, which was used without any additional purification in the next step.

N*-[(*tert*-Butoxy)carbonyl]-3-[(1*H*-indol-3-ylmethylene)amino]-*L*-tyrosine Methyl Ester* (**3**). To a soln. of **2** (4 mmol) in abs. EtOH (6 ml) was added a soln. of 1*H*-indole-3-carboxaldehyde (4 mmol, 0.58 g) in abs. EtOH (6 ml). The mixture was boiled for 5 min, cooled, and stirred at r.t. overnight (TLC monitoring (AcOEt/petroleum ether 2 : 5): R_f 0.61; no substrate left). After this time, the solvent was evaporated and the product, without any additional purification, was used at once in the next step.

N*-[(*tert*-Butoxy)carbonyl]-3-[2-(1*H*-indol-3-yl)benzoxazol-5-yl]-*L*-alanine Methyl Ester* (**4**). a) *Cyclization of* **3** *in the Presence of Lead Tetraacetate*. To a soln. of **3** (1 mmol, 437 mg) in AcOH (10 ml), lead tetraacetate (1 mmol, 443 mg) was added at r.t. After a few minutes, the solvent was evaporated, the oily residue treated with AcOEt, and the soln. washed with H_2O (3 × 5 ml), 0.1M $KHSO_4$ (3 × 5 ml), and a sat. NaCl soln. (3 × 5 ml), dried ($MgSO_4$) and evaporated: crude **4**.

b) *Cyclization of* **3** *in the Presence of NBS*. To a soln. of **3** (3.84 mmol, 1.68 g) in CH_2Cl_2 (10 ml) cooled to 0° in an ice bath, NBS (4.23 mmol, 0.75 g) was added, and the mixture was stirred for *ca*. 30 min. Then the solvent was evaporated, the brown, oily residue dissolved in AcOEt (30 ml), and the org. phase washed with H_2O (2 × 10 ml), 5% $NaHCO_3$ soln. (2 × 10 ml), and sat. NaCl soln. (2 × 10 ml), dried ($MgSO_4$), and evaporated: crude **4/5**.

c) *Isolation of* **4** *and* **5**. Both products **4** and **5** were isolated from the crude mixture **4/5** or from the crude **4** semi-prep. reversed-phase HPLC (24 → 80% MeCN/H_2O (with addition of 0.1% CF_3COOH soln.) in 120 min). Recrystallization from MeCN gave crystalline, **4** and **5**. The total yield of **4**, regardless of the cyclization reagent used, was *ca*. 7.2% (125.9 mg), and that of **5** 3.6% (74.2 mg). The purity of **4** and **5** was checked by anal. reversed-phase HPLC. (0 → 80% MeCN/H_2O (with addition of 0.1% CF_3COOH soln.) in 60 min): t_R 50.61 (**4**) and 57.24 min (**5**).

d) *Data of* **4**: Light yellow solid. M.p. 201–202°. $[\alpha]_D^{20} = +11.68$ ($c = 0.01$M, MeOH). FT-IR: 3352.7*s*, 1742.5*s*, 1714.1*s*, 1626.3*s*, 1574.0*s*, 1436.9*m*, 1391.8*w*. ^1H-NMR: 1.44 (*s*, ^tBu); 3.24 (*m*, $CH_2(\beta)$); 3.75 (*s*, MeO);

4.67 (*d*, H–C(α)); 5.03 (*d*, NH); 7.07 (*d*, H–C(6)); 7.36 (*m*, H–C(5′), H–C(4)); 7.48 (*m*, H–C(6′), H–C(7)); 7.51 (*d*, H–C(7′)); 8.09 (*d*, H–C(2′)); 8.49 (*m*, H–C(4′)); 8.63 (*d*, H–N(1′)). ^{13}C-NMR: 28.53 (Me_3C); 38.62 (C(β)); 52.50 (MeO); 54.98 (C(α)); 80.82 (Me_3); 105.75 (C(3′)); 110.10 (C(7)); 111.79 (C(6′)); 119.96 (C(7′)); 121.82 (C(4′)); 122.23 (C(5′)); 123.82 (C(4)); 125.26 (C(6)); 125.33 (C(7′a)); 127.64 (C(2′)); 132.33 (C(3′a)); 136.51 (C(5)); 143.21 (C(3a)); 149.41 (C(7a)); 151.83 (C(2)); 162.03 (HNCOO); 172.53 (C(α)*C*OO). Anal. calc. for ($C_{24}H_{25}N_3O_5$): C 66.19, N 9.65, H 5.79; found: C 66.14, N 9.58, H 5.73. ^{1}H,^{13}C-HSQC: 8 cross-peaks within 7–9 ppm (arom. H) and 100–160 ppm (arom. C). FAB-MS: 436 ([$M+1$]$^+$).

e) *Data of 3-[4-Bromo-2-(1*H*-indol-3-yl)benzoxazol-5-yl]-*N*-[(*tert*-butoxy)carbonyl]-*L*-alanine Methyl Ester* (**5**): White solid. M.p. 203–204°. $[\alpha]_D^{20} = +2.35$ ($c = 0.01$M, MeOH). FT-IR: 3416.0*s*, 3340.3*s*, 1732.7*s*, 1687.4*s*, 1630.1*m*, 1603.6*m*, 1578.8*s*, 1522.1*m*, 1057.0*m*. ^{1}H-NMR: 1.44 (*s*, ^{t}Bu); 3.23 (*m*, CH_2(β)); 3.77 (*s*, MeO); 4.65 (*d*, H–C(α)); 5.06 (*d*, NH); 7.22 (*d*, H–C(6)); 7.37 (*m*, H–C(6′), H–C(5′)); 7.43 (*d*, H–C(7)); 7.49 (*m*, H–C(7′)); 8.14 (*d*, H–C(2′)); 8.48 (*m*, H–C(4′)); 8.71 (*d*, H–N(1′)). ^{13}C-NMR: 28.53 (*Me*$_3$C); 38.26 (C(β)); 52.65 (MeO); 54.86 (C(α)); 80.43 (Me_3C); 101.99 (C(3′)); 105.07 (C(4)); 111.86 (C(7′)); 118.95 (C(7)); 121.78 (C(4′)); 122.45 (C(6′)); 123.97 (C(5′)); 125.17 (C(7′a)); 128.20 (C(2′)); 128.29 (C(6)); 134.01 (C(3′a)); 136.45 (C(5)); 143.64 (C(3a)); 147.55 (C(7a)); 155.43 (C(2)); 161.84 (HNCOO); 172.25 (C(α)*C*OO). ^{1}H,^{13}C-HSQC: 7 cross-peaks within 7–9 ppm (arom. H) and 100–160 ppm (arom. C). ^{1}H,^{1}H TOCSY: 3 spin systems, *i.e.* H–C(4′)/H–C(5′)/H–C(6′)/H–C(7′) (indole moiety) H–N(1′)/H–C(2′) (indole moiety), and H–C(6)/H–C(7) (benzoxazole moiety). FAB-MS: 514 and 516 ([$M+1$]$^+$).

REFERENCES

[1] J. R. Lakowicz, 'Principles of Fluorescence Spectroscopy', 2nd edn., Kluwer Academic/Plenum Publishers, New York, Boston, Dordrecht, London, Moscow, 1999.

[2] A. G. Szabo, D. M. Rayner, *J. Am. Chem. Soc*. **1980**, *102*, 554.

[3] R. F. Chen, J. R. Knutson, H. Ziffer, D. Porter, *Biochemistry* **1991**, *30*, 5184.

[4] J. B. A. Ross, H. R. Wyssbrod, R. A. Porter, G. P. Schwartz, C. A. Michaels, W. R. Laws, *Biochemistry* **1992**, *31*, 1585.

[5] M. Beechem, L. Brand, *Ann. Rev. Biochem*. **1985**, *54*, 43.

[6] M. R. Eftink, *Methods Biochim. Anal*. **1991**, *35*, 127.

[7] Z. Bajzer, F. G. Prendergast, *Biophys. J.* **1993**, *65*, 2313.

[8] A. Rzeska, K. Stachowiak, J. Malicka, L. Łankiewicz, W. Wiczk, *J. Photochem. Photobiol. A: Chem*. **2000**, *133*, 33.

[9] W. Wiczk, K. Stachowiak, P. Skurski, L. Łankiewicz, A. Michniewicz, A. Rój, *J. Am. Chem. Soc*. **1996**, *118*, 8300.

[10] M. L. McLaughlin, M. D. Barkley, *Methods Enzymol*. **1997**, *278*, 190.

[11] M. R. Eftink, Y. Jia, D. Hu, C. A. Ghiron, *J. Phys. Chem*. **1995**, *99*, 5713.

[12] J. B. A. Ross, A. G. Szabo, C. W. Hogue, *Methods Enzymol*. **1997**, *278*, 151.

[13] W. R. Laws, G. P. Schwartz, E. Rusinva, G. T. Burke, Y. C. Chu, P. G. Katsoyannis, J. B. A. Ross, *J. Protein Chem*. **1995**, *14*, 225.

[14] C. W. V. Hogue, A. G. Szabo, *Biophys. Chem*. **1993**, *48*, 159.

[15] R. L. Rich, M. Nergerie, J. Li, S. Eliott, R. W. Thornburg, J. W. Petrich, *Photochem. Photobiol*. **1993**, *58*, 28.

[16] A. V. Smirnov, D. S. English, R. L. Rich, J. Lane, L. Teyton, A. W. Schwabacher, S. Luo, R. W. Thornburg, J. W. Petrich, *J. Phys. Chem. B* **1997**, *101*, 2758.

[17] R. W. Thornburg, J. W. Petrich, *J. Phys. Chem. B* **1997**, *101*, 2758.

[18] C. Y. Wong, M. R. Eftink, *Biochemistry* **1998**, *37*, 8938.

[19] D. S. English, R. L. Rich, J. W. Petrich, *Photochem. Photobiol*. **1998**, *67*, 76.

[20] R. P. Haugland, ' Handbook of Fluorescent Probes and Research Chemicals', 6th edn., Molecular Probes, 1996.

[21] M. Koller, H. Eckert, *Anal. Chim. Acta* **1997**, *352*, 31.

[22] Y. Kanegae, K. Peariso, S. S. Martinez, *Appl. Spectrosc.* **1996**, *50*, 316.

[23] B. M. Krasovitskii, B. M. Bolotin, 'Organic Luminescent Materials', VCH, Weinheim, 1988, p. 169.

[24] A. Pla-Dalmau, *J. Org. Chem*. **1995**, *60*, 5468.

[25] B. M. Krasovitskii, B. M. Bolotin, 'Organic Luminescent Materials', VCH, Weinheim, 1988, p. 89.

[26] Y.-H. So, J. P. Heeschen, *J. Org. Chem*. **1997**, *62*, 3552.

[27] M. R. DeLuca, S. M. Kerwin, *Tetrahedron* **1997**, *53*, 457.

[28] M. R. DeLuca, I. B. Tarapolewala, S. M. Kerwin, *Heterocycles* **1999**, *51*, 979.
[29] F. F. Stephens, J. D. Bower, *J. Chem. Soc*. **1949**, 2971.
[30] F. F. Stephens, J. D. Bower, *J. Chem. Soc*. **1950**, 1722.
[31] R. G. Srivastava, P. S. Venkataramani, *Synth. Commun.* **1988**, *18*, 1537.
[32] E. Wünsch, 'Synthese von Peptiden', ' Houben-Weyl, Methoden der organischen Chemie', Ed. E. Müller, Georg Thieme Verlag, Stuttgart 1974, Vol. 15, p. 317.
[33] J. M. Stewart, J. D. Young, 'Solid Phase Peptide Synthesis', 2nd edn., Pierce Chemical Company, Rockford, Illinois 1984, p. 63.
[34] R. F. Chen, *Anal. Lett*. **1967**, *1*, 35.
[35] D. F. Eaton, *Pure Appl. Chem*. **1988**, *60*, 1107.

Received January 30, 2001